Intelligibilité du numérique

N° 7 | 2025
Le 4ème paradigme en question

Comment fait-on parler les données par elles-mêmes ? Une approche communicationnelle de la production de connaissance

Cassandre BURNIER *Docteur*
Information et Communication
ReSIC
Université Libre de Bruxelles, Brussels

Édition électronique :
URL :
https://intelligibilite-numerique.numerev.com/articles/revue-7/3214-comment-fait-on-parler-les-donnees-par-elles-memes-une-approche-communicationnelle-de-la-production-de-connaissance
Date de publication : 04/05/2025

CertiScience© *Certifié évalué par les pairs*

Cette publication est sous licence **CC BY-NC-ND** (Attribution - No commercial - No derivatives).

Pour **citer cette publication** : BURNIER, C. (2025) Comment fait-on parler les données par elles-mêmes ? Une approche communicationnelle de la production de connaissance. *Intelligibilité du numérique*, (7). https://doi.org/10.34745/

Cet article déconstruit le mythe de la donnée « auto-explicative » en explorant les processus par lesquels elle acquiert du sens et une capacité d'action. En articulant la théorie de la ventriloquie, l'agencement narratif et la rhétorique de l'intelligibilité, l'étude démontre que les données ne s’expriment jamais seules : leur agentivité résulte d'un travail collectif de mise en récit et de qualification. L’enquête ethnographique révèle que l’intelligibilité des données est un construit sociotechnique, négocié en permanence entre les logiques techniques et les impératifs métiers. Le cas du « Profil SSO », loin d'être une mesure neutre, apparaît comme une biographie numérique de l’utilisateur façonnée par des interactions humaines et algorithmiques. Finalement, ce travail propose une approche communicationnelle de la production de connaissance pour répondre aux enjeux du quatrième paradigme.

How do we make data speak for itself? A communicational approach to knowledge production

This article deconstructs the myth of "self-explanatory" data by exploring the processes through which it acquires meaning and the capacity for action. By articulating the theory of ventriloquism, narrative assembly (agencement), and the rhetoric of intelligibility, the study demonstrates that data never speaks alone: its agency results from a collective effort of storytelling and qualification. The ethnographic inquiry reveals that data intelligibility is a sociotechnical construct, constantly negotiated between technical logics and business imperatives. The "SSO Profile" case, far from being a neutral metric, emerges as a digital biography of the user shaped by human and algorithmic interactions. Ultimately, this work proposes a communicational approach to knowledge production to address the challenges of the "fourth paradigm".

Mots-clés :
Ethnographie, Approche constitutive, Données massives, Ventriloquie, Agencement narratif, Big data, Ethnography, Constitutive approach, Ventriloquism, Narrative arrangement

1. Introduction

Au cœur des organisations contemporaines, et singulièrement au sein des médias, les données massives se sont imposées comme des ressources stratégiques majeures. Présentées comme des indicateurs neutres et fiables, elles sont investies d'une mission de rationalisation consistant à guider les décisions et objectiver les processus. Cette foi dans leur objectivité assure leur légitimité dans les espaces décisionnels (Kitchin, 2014; van Dijck, 2014) et s'inscrit dans le cadre du quatrième paradigme scientifique (Hey et al., 2009), lequel érige l'exploration algorithmique en méthode de découverte souveraine[1].

Ce paradigme repose sur une idée radicale : les traces numériques seraient capables de parler d'elles-mêmes. Leur simple mise en relation automatisée suffirait à produire du sens, sans recours nécessaire à une théorie ou à un récit préexistant. Cette vision nominaliste (Bachimont, 2014, 2015) tend à occulter les opérations humaines, techniques et organisationnelles qui rendent ces chiffres interprétables. Or, comme le soulignent de nombreux auteurs(boyd & Crawford, 2012; Couldry, 2017; Gitelman, 2013; Kitchin, 2014; Mayer-Schönberger & Cukier, 2013), les données ne sont jamais brutes : elles sont toujours *déjà* le produit de choix, de conventions et de dispositifs de captation.

Dans le secteur médiatique, cette tension est omniprésente. Les mesures d'audience orientent les choix éditoriaux, redéfinissent les formats et influencent les narrations. Pourtant, leur mise en sens repose sur une série d'interactions *situées* entre acteurs et artefacts : ingénieurs, analystes, outils de mesure et architectures informatiques. Ces agencements traduisent des arbitrages, des négociations et des rapports de force (Ananny & Crawford, 2016; Burnier, 2023).

Cet article propose une approche communicationnelle de la production de connaissance par les données. Il postule que ces dernières ne s'expriment jamais seules, mais sont rendues intelligibles par des processus d'agencement narratif et de ventriloquie impliquant une pluralité d'intervenants. En mobilisant les théories de la Communication Constitutive des Organisations (CCO) et de la rhétorique de l'intelligibilité (Bachimont, 2020), il s'agit d'analyser comment ces ressources acquièrent leur autorité au sein des pratiques professionnelles.

À partir d'une ethnographie menée dans un groupe de presse belge, l'article examine la

construction du « Profil SSO » comme *biographie numérique de l'utilisateur*. Ce dispositif de *Single Sign-On* (identification unique), croisé avec un progiciel de gestion d'abonnements (un ERP, tel que Compiere) et des outils de mesure d'audience (tels que Google Analytics), alimente une plateforme en nuage (Microsoft Azure) pour générer des profils exploitables à des fins marketing et éditoriales. C'est cette chaîne d'intégration, dont l'articulation technique est détaillée ultérieurement dans l'analyse, qui constitue notre terrain de recherche. Ce cas démontre que l'intelligibilité des données est le fruit d'un travail collectif, situé et souvent invisible, où l'humain et la technique co-construisent un récit pour faire parler les chiffres, parfois malgré eux.

2. La donnée comme actant : ventriloquie, agencement et intelligibilité

Dans le sillage du quatrième paradigme scientifique (Hey et al., 2009) où les données massives sont perçues comme des entités auto-explicatives, cette recherche interroge les mécanismes par lesquels ces traces acquièrent du sens et une capacité d'action. L'approche communicationnelle mobilisée ici articule trois perspectives : la rhétorique de l'intelligibilité, la ventriloquie et l'agencement narratif. Ces théories convergent pour démontrer que le sens émerge des interactions sociotechniques plutôt que d'être une propriété immanente aux chiffres.

L'approche constitutive de la communication organisationnelle (CCO) postule que les organisations ne préexistent pas à la communication, mais sont continuellement constituées par les processus communicationnels (Taylor, 1988). Dans cette perspective, les artefacts techniques ne sont pas de simples vecteurs d'information, mais des éléments performatifs façonnant la réalité. Les données sont ainsi considérées comme des actants participant activement à la constitution de l'organisation (Bencherki et al., 2023). Cette vision s'appuie sur la théorie du sensemaking de Weick (1995) concevant le sens comme une fabrication rétrospective où l'organisation est un flux d'interprétations stabilisées. L'aphorisme « l'oignon n'a pas de noyau » (Taylor et al., 1996) illustre parfaitement cette conception d'une structure émergente, sans essence fixe, toujours en devenir.

Cette ontologie est prolongée par la théorie de la ventriloquie, qui offre un outil conceptuel pour penser l'agentivité distribuée entre humains et autres- qu'humains (Cooren, 2010, 2013). Elle démontre que la capacité d'agir est *coproduite* : un ventriloque humain peut faire parler une marionnette technique, mais l'inverse est tout aussi vrai. Les données peuvent être ventriloquées pour justifier des choix stratégiques, tout comme elles peuvent ventriloquer les acteurs en imposant des logiques de performance. L'agentivité n'est donc pas une propriété intrinsèque, mais un effet relationnel produit dans *l'ici et maintenant* de l'interaction (Cooren & Robichaud, 2006). Ce mécanisme de porte-parolat transforme les données en figures d'autorité capables d'orienter les décisions (Burnier, 2023).

Toutefois, si la ventriloquie explique comment l'on fait parler les données, elle ne précise pas les conditions rendant cette parole intelligible. C'est ici qu'intervient l'agencement narratif, conçu comme une configuration sociomatérielle où données, scripts et discours s'articulent pour produire des récits opérants (Gherardi, 2016; Marion, 1997; Ricoeur, 2014). Puisque les corrélations ne peuvent générer leur propre causalité (Bachimont, 2020), l'humain doit les insérer dans un agencement qui les rend racontables. Cette perspective rejoint l'ontologie narrative de Schapp, selon laquelle les individus sont « empêtrés dans des histoires » qui configurent leur monde (Rodrigues, 2019). L'intelligibilité devient alors le résultat d'un processus rhétorique et technique qui rend les données aptes à circuler et à faire agir.

L’agencement narratif renvoie aux conditions de possibilité de l’intelligibilité ou aux conditions idéales de production (Cooren, 2013). Ces dernières ne sont pas des règles rigides, mais des modèles itérables qui se stabilisent par la répétition. Cette itérabilité assure au récit sa double fonction : la narration des faits peut être répétée pour circuler dans l'organisation et les types de situations peuvent se reproduire. Le schème narratif, une fois stabilisé, devient un vecteur d'intelligibilité généralisable. Dans le cas du « Profil SSO » que nous verrons, les règles élaborées pour résoudre une anomalie deviennent des schèmes réutilisables pour des problèmes similaires. L'agencement narratif constitue ainsi l'opération énonciative instaurant le sens dans le discours (Schulz, 2017), privilégiant la plausibilité et la cohérence (Weick, 1995) sur la simple véracité.

En articulant CCO, ventriloquie et rhétorique de l'intelligibilité, ce cadre propose une compréhension nuancée de la production de connaissances. Il résout la tension entre l'agentivité des données et leur incapacité à générer leur propre causalité en montrant que leur parole est le fruit d'un travail collectif. Cette perspective rejette le déterminisme du quatrième paradigme pour insister sur *la nécessité du récit*. Elle déplace l'analyse des données comme objets passifs vers une vision de protagonistes actifs au sein de récits coconstitués. Tandis que la CCO fournit le cadre ontologique et la ventriloquie la grille analytique, la rhétorique de l'intelligibilité pose la question épistémique des conditions de compréhension, offrant ainsi une vision intégrale des dynamiques sociotechniques contemporaines.

3. Une ethnographie de l’intelligibilité des données

Cette recherche adopte une posture qualitative visant à comprendre comment les flux d’information acquièrent du sens et une capacité d’action au sein des organisations médiatiques. Sur le plan méthodologique, la démarche relève d'une ethnographie organisationnelle à visée compréhensive, permettant d’observer le travail des données en train de se faire et de saisir la complexité des processus de production de sens dans leur contexte quotidien (Rouleau, 2013; Van Maanen, 2011). L’étude s’inscrit dans le

cadre d’une immersion prolongée de près de trois ans au sein de deux groupes de presse belges francophones, Rossel et IPM. L’accès à ce terrain, fruit de négociations approfondies, a garanti une transparence totale quant à mon rôle de chercheure. Afin de préserver l’intégrité des acteurs, les noms des individus et de certaines entités internes ont été systématiquement anonymisés.

L’observation prolongée, les entretiens et la vidéo-ethnographie constituent les principales techniques de collecte. Bien que ma recherche doctorale articule de manière itérative plusieurs méthodes (filature, entretien, observation), le présent article se concentre sur une étude de cas unique. Le corpus sélectionné comprend un entretien de 54 minutes avec une ingénieure de la donnée, une vidéo-ethnographie de 8 minutes capturant ses gestes techniques lors du traitement d’une table de profil utilisateur dans l’outil Microsoft Azure, ainsi que des notes ethnographiques issues d’échanges informels. L'approche de la ventriloquie intervient ici comme cadre d'analyse des matériaux empiriques : elle guide le codage et l'interprétation des interactions observées en identifiant les dynamiques par lesquelles les acteurs humains et les artefacts techniques se font mutuellement parler. Le choix de ce cas est motivé par sa capacité à révéler les anomalies des traces numériques liées au « Profil SSO » (*Single Sign-On*) et la nécessité de solliciter les départements métiers pour définir la vérité de la donnée[2].

L'analyse des matériaux empiriques a été réalisée de manière itérative, selon un mouvement de va-et-vient entre le travail de terrain (*fieldwork*), la conceptualisation réflexive (*headwork*) et l'effort de rédaction (*textwork*) préconisé par Van Maanen (2011). Les verbatims et les séquences vidéo ont fait l’objet d’un codage thématique visant à faire émerger les notions clés et les logiques d’interaction. Bien que les données du « Profil SSO » ne constituent pas un échantillon massif au sens strictement volumétrique, elles répondent aux caractéristiques fondamentales des données massives définies par Kitchin (2014) : une génération automatique et continue, une provenance de sources hétérogènes nécessitant une intégration complexe et un traitement via des infrastructures de cloud computing. C’est précisément cette complexité relationnelle, plutôt que le volume absolu, qui rend nécessaire le travail d’intelligibilité analysé ici.

Cette approche permet de dépasser une vision purement instrumentale pour révéler les dimensions narratives, performatives et politiques de la donnée. Elle éclaire la manière dont ces ressources sont agencées dans des contextes organisationnels marqués par des tensions entre automatisation, pluralité des récits et rationalisation algorithmique. Cette démarche s’inscrit enfin dans une posture réflexive située au sein d’un cadre épistémologique occidental et de ses normes méthodologiques dominantes.

4. Faire parler le Profil SSO, raconter la biographie

de l’utilisateur

L’analyse se concentre sur un point névralgique de l’ethnographie : la co-constitution itérative de l’identité des usagers à travers le développement du « Profil SSO ». Ce système d’authentification unique centralise la gestion des identités numériques, permettant de lier les accès aux différents services du groupe de presse (portail web, édition numérique, espace abonné) sous un identifiant souverain. Plus précisément, l’étude porte sur la construction d’une table de dimension au sein de l’infrastructure Microsoft Azure. Cette plateforme d’informatique en nuage (*cloud computing*) agrège et traite les flux nécessaires à la gestion des abonnements, issus d’un progiciel ERP (Compiere), et les traces comportementales collectées via Google Analytics. L’objectif est de stabiliser ces sources hétérogènes pour offrir une vue unifiée du parcours client. Pour Alice, l’ingénieure chargée du projet, cette table doit permettre de transformer des lignes de codes en récits de vie : il s’agit de savoir qu’un utilisateur spécifique a consulté des articles de sport durant six minutes tout en possédant un abonnement actif à une publication précise.

Cependant, l’ingénieure se heurte rapidement aux limites de la qualité et de la cohérence des matériaux bruts. Dans un système idéal, chaque abonné correspondrait à un identifiant unique, une adresse email et un contrat. Or, l’historique des modes de collecte a transformé la base de données en un labyrinthe technique : un même identifiant peut pointer vers plusieurs adresses, tandis qu’un individu unique peut se voir attribuer plusieurs codes SSO distincts. Alice décrit cette réalité lors d’un entretien en mars 2020 :

> *Alice – Du coup, il faut reprendre la table XXSSO et comprendre quels sont tous les cas qu’on peut avoir dans cette table-là. Donc, on peut avoir le cas où effectivement, tu as un SSO unique pour un certain tiers, pour une certaine organisation, pour une certaine adresse email. En fait, tu as les cas aussi où tu as le même SSO, il est en deux lignes parce que l’adresse email est différente. Ou le même SSO est en trois lignes parce que c’est le même SSO mais il est relié à trois tiers différents. Tous des trucs comme ça. (Entretien, mars 2020)*

Ces anomalies décrites par Alice – un même SSO lié à plusieurs adresses email, plusieurs SSO pour une même personne – agissent comme des agents perturbateurs. Elles résistent à l’intelligibilité immédiate et à l’agrégation automatisée, forçant l’ingénieure à s’engager dans un travail de production de sens avec ses artefacts. Cette résistance du matériau est une manifestation de son agentivité : la trace numérique fait faire des choses aux humains, les obligeant à entrer dans un processus de négociation. L'incohérence démontre que le chiffre ne parle jamais d'un ton univoque ; au contraire, il soulève des questions sur sa propre représentativité.

Dans la perspective de l’agencement narratif, les discordances décrites par Alice constituent l’intrigue au sens de Ricœur (2014) : un événement qui rompt l'ordre

attendu et appelle une explication. La mise en intrigue s'opère lorsque l'ingénieure catégorise et énumère systématiquement ces cas ("tu as le cas où...", "les cas aussi où..."), transformant un désordre brut en une série de configurations identifiables. Ce travail révèle que la donnée ne porte pas en elle-même la frontière entre le normal et le pathologique. Cette indétermination oriente alors l'enquête vers une nécessité de médiation : la consultation des départements métiers pour trancher la vérité de la donnée.

Face à ces incohérences, l'ingénieure de la donnée ne peut décider seule des règles de traitement, comme l'exclusion systématique des identifiants liés à plusieurs adresses. Elle doit impérativement consulter le *Business* pour déterminer si ces anomalies sont normales d'un point de vue métier. Ce processus de validation est décrit par Alice lors de notre entretien en mars 2020 :

> *Alice – Et donc comprendre et après, aller essayer avec le Business pour voir : « Est-ce que c'est normal ? Est-ce que tu peux avec ton ID SSO... Est-ce qu'il peut être relié à plusieurs personnes différentes ? Est-ce que ton ID SSO, il peut être relié à plusieurs mails différents... Est-ce qu'en fait, c'est l'ingestion de données qui a été mal faite et donc, il faut exclure ces cas ? ». Pour comprendre ce qui se cache derrière ces données. Parce que nous on ne peut pas, fin je ne peux pas prendre des décisions sur tous ces cas-là : « je les ai exclus et ils ne vont jamais se retrouver dans ma dimension et donc, du coup ils ne vont jamais pouvoir se faire liés avec du trafic Google Analytics ». Parce que peut-être qu'en fait, c'est normal qu'un certain ID SSO ait plusieurs adresses mails parce qu'ils autorisent peut-être d'avoir une adresse mail primaire, une adresse mail secondaire. Si c'est normal, alors c'est à nous de réfléchir : comment est-ce que je vais faire pour que tout soit unique dans mes données et avec quoi je vais les lier ? (Entretien, mars 2020)*

Ce dialogue révèle que la vérité de la donnée n'est pas un fait technique objectif, mais un construit organisationnel. La donnée SSO, par sa plasticité et ses ambiguïtés, agit comme un agent qui force la renégociation des conventions de mesure. L'enquête prend alors la forme d'un dialogue ventriloque : Alice fait parler la donnée devant le *Business* en l'interrogeant sur ses propriétés possibles, tandis que les responsables métier, en répondant, font parler leurs propres logiques à travers elle. L'agentivité est ici *distribuée* : la trace numérique, par ses caractéristiques intrinsèques, initie la discussion, mais son sens final est le fruit d'une décision collective.

La narration résolutive, qui constitue l'intelligibilité finale, émerge de cette négociation. Elle prend la forme d'une règle de traitement technique (par exemple, la hiérarchisation entre adresse primaire et secondaire) qui transforme l'anomalie initiale en *un récit cohérent de l'identité de l'utilisateur*. En stabilisant ainsi *ce qui fait vérité*, les acteurs transforment un signal bruyant en une biographie numérique exploitable. Ce processus démontre que rendre les données intelligibles ne revient pas à découvrir un sens préexistant, mais à produire un agencement narratif capable de réconcilier les contraintes du système technique et les besoins du métier.

La construction du « Profil SSO » est un processus itératif, rythmé par des cycles de développement technique et de validation métier. Alice explique que le produit fini n'est jamais livré de manière monolithique, mais se construit par sédimentation. Ce déploiement par étapes révèle que l'agencement narratif n'est pas une opération ponctuelle, mais un récit feuilletonné. Chaque épisode – qu'il s'agisse du choix des champs, de la validation des formats ou du traitement de l'historique – constitue une mise en intrigue partielle qui s'agrège progressivement pour former la biographie numérique de l'utilisateur. Lors de notre entretien, l'ingénieure souligne cette progression :

> *Alice – Souvent, les développements se font par mini-étapes. Donc là, souvent il y a une question qui est : Qu'est-ce qu'on fait au niveau de l'historique ? Donc là, par exemple, de nouveau avec la Dim SSO [sourire crispé], on n'y a jamais touché. (...) Et donc, si on commence à la traiter et à sortir des données, bah on va sortir les données à la date d'aujourd'hui. (...) Donc, ça veut dire que toutes mes données des abonnements avant aujourd'hui, je ne pourrai jamais savoir quel est leur SSO si je ne fais pas de traitements supplémentaires. (...) Mais ce genre de traitements, on ne le fait pas tant qu'on n'a pas l'accord du Business. (...) Et ça, du coup, tu ne peux pas prévoir. Tu ne le prévois pas dans ton développement et tu avances étape par étape. Tu prévois : Comment est-ce que tu vas traiter cette table au niveau Business ? Je vais sortir l'email, je vais sortir ça et ça. Après, ils confirment : Est-ce que c'est bien les champs qu'ils voulaient ou non ? Ensuite, ils confirment s'ils ont besoin de l'historique, oui ou non ? Ensuite, ils confirment si avec l'historique tu peux faire une image comme aujourd'hui et la mettre avant ou si tu dois aller rechercher dans tous les anciens records [enregistrements] de la table XXSSO comment elle était chaque jour. Le Business, c'est étape par étape. On ne délivre jamais un produit fini. (Entretien, mars 2020)*

Chaque étape de ce développement est une mise en intrigue où des choix techniques et métier façonnent la trajectoire de l'usager. La décision d'inclure ou d'exclure l'historique des données est un exemple frappant : elle *détermine la profondeur temporelle* de la biographie que la donnée pourra raconter. Ainsi, le rôle de l'ingénieure dépasse largement l'exécution technique pour devenir profondément communicationnel. Elle mobilise des ressources rhétoriques pour structurer les données de manière à les rendre logiques et légitimes, quitte à simplifier ou normaliser certains éléments pour les faire entrer dans le cadre narratif attendu.

Cette traduction constante entre contraintes techniques et besoins métier n'est pas sans heurts. Le sourire crispé d'Alice lorsqu'elle évoque les demandes du *Business* témoigne d'une tension inhérente à la négociation des récits. La donnée ne devient intelligible que lorsque son agencement technique s'aligne avec les nécessités pragmatiques de l'organisation. En définitive, le « Profil SSO » émerge comme une construction sociotechnique où l'ingénieure et le *Business* collaborent pour fabriquer une identité numérique actionnable.

> *Alice – Le seul moyen de connecter ta base de données "abonnement" qui est chez toi, qui est un ERP, avec ce qui se passe sur ton site internet, c'est si sur ton site tu récupères ce petit ID et du coup, tu as ton trafic du site relié à un ID, tes abonnements reliés à un ID. Tu les connectes ensemble et tu es content. Tu peux savoir que le Monsieur qui a passé six minutes sur les articles sport, il a un abonnement Le Soir web je ne sais pas quoi. Là, ce n'est pas trop possible, il y a plein de problème comme ça. (Entretien, mars 2020)*

Cette finalité illustre l'aboutissement de la co-construction : permettre à la donnée de *raconter une histoire spécifique*, comme celle de cet amateur de sport abonné à la version numérique ("le Monsieur qui a passé six minutes sur les articles sport et qui a tel abonnement"). L'agentivité de la donnée (sa capacité à influencer les stratégies marketing ou éditoriales) dépend directement de la réussite de cet agencement narratif. L'intelligibilité n'est donc pas une caractéristique intrinsèque de l'information, mais le produit d'un travail collectif continu où humains et systèmes techniques co-écrivent le sens des traces numériques.

5. Discussion

Cet article visait à déconstruire le mythe de la donnée auto-explicative en explorant les mécanismes par lesquels les données massives acquièrent sens et capacité d'action dans les organisations, en particulier au sein des médias. L'étude ethnographique de la construction du « Profil SSO » au sein d'un groupe de presse a démontré que l'intelligibilité des données est loin d'être un fait technique neutre. Elle est un construit social et organisationnel, le produit d'un travail collectif, situé et itératif, où les humains et les systèmes techniques s'influencent mutuellement pour faire parler les données.

Les résultats démontrent que l'identité de l'utilisateur, telle que construite à travers le Profil SSO, est loin d'être une représentation neutre ou intrinsèquement objective. Au contraire, elle est le produit d'une co-constitution itérative et d'un agencement narratif continu, où les données dites « brutes », souvent incohérentes quand elles ne se parlent pas entre elles ou ne racontent pas une histoire *plausible(Weick, 1995, 2020)*, agissent comme des agents perturbateurs, forçant les humains à engager un travail de production de sens. L'ingénieure de la donnée, confrontée aux ambiguïtés – un même SSO lié à plusieurs emails ou un abonnement avec plusieurs SSO – ne peut décréter seule *la vérité de la donnée*. Elle doit impérativement consulter d'autres acteurs de l'organisation – le *Business*, les données, l'ERP, *Microsoft Azure*, le principe de centralisation –, révélant ainsi que la vérité de la donnée est un construit social et organisationnel, négocié entre les logiques techniques et les réalités métier (Colletis et al., 2020; Leonelli, 2014).

Cette négociation illustre la rhétorique de l'intelligibilité (Bachimont, 2020) selon

laquelle le travail humain est essentiel pour fournir une explication causale et une profondeur sémantique aux corrélations brutes fournies par la machine. « Le storytelling est l'ersatz de l'intelligibilité » lorsque le calcul ne suffit pas à expliquer par lui-même, soulignant le rôle crucial de la narration pour donner du sens aux résultats des manipulations calculatoires (Bachimont, 2020). La ventriloquie (Cooren, 2010, 2013) est le concept idéal pour décrire cette interaction : la donnée (la marionnette) « parle » parce que les humains (les ventriloques) la font parler, mais aussi, les caractéristiques de la donnée (ses anomalies, sa résistance) peuvent « faire parler » ou agir les humains de certaines manières. Cette dynamique relationnelle est la clé pour comprendre la donnée comme *un protagoniste* sans lui attribuer une capacité d'auto-génération de sens.

L'approche de la Communication Constitutive des Organisations (CCO) montre concrètement que l'organisation, centrée ici sur l'identité de ses utilisateurs et utilisatrices, n'est pas une entité préexistante, mais qu'elle est continuellement façonnée par les interactions entre une pluralité d'acteurs (Schoeneborn et al., 2025; Taylor et al., 1996). La donnée, par sa capacité à « faire une différence » (Cooren, 2013, p. 112), agit comme un agent qui interpelle et initie le processus de production de sens. L'ingénieure de la donnée, en interrogeant les incohérences du SSO, ventriloque la donnée en lui posant des questions au nom du système technique. Inversement, le *Business* ventriloque la donnée en lui attribuant une normalité ou une anormalité basée sur ses propres logiques métier. Ainsi, l'organisation médiatique naît d'une oscillation constante entre les rôles de ventriloque et de marionnette, que les humains et les choses prennent tour à tour : les individus font parler les choses et d'autres individus, tout autant que ces choses et ces individus les font parler par leurs paroles ou leurs actes (Cooren, 2010; Nathues et al., 2020). L'analyse ventriloque révèle la multiplicité des voix à travers le discours en interaction : la donnée initie la discussion, mais son sens final est le fruit d'une décision collective, reflétant une ontologie relationnelle où l'agence est partagée entre humains et autres-qu'humains (Nathues et al., 2020; Rodrigues, 2019).

L'analyse du processus itératif de construction du Profil SSO met en lumière l'importance de l'agencement narratif (Bachimont, 2020; Gherardi, 2016; Rodrigues, 2019). Les données brutes, disparates et ambiguës, sont progressivement agencées – nettoyées, transformées, liées – pour raconter une histoire cohérente de l'utilisateur. Chaque étape de développement, chaque décision sur l'inclusion de l'historique ou le traitement des anomalies, constitue une « mise en intrigue » (Ricoeur, 2014) qui façonne le récit final que la donnée pourra porter. Pour que les données du Profil SSO puissent "raconter une histoire" de l'utilisateur (par exemple, "ce Monsieur a passé six minutes sur les articles sport et a tel abonnement"), il faut qu'il existe des « conditions idéales de production » (Cooren, 2013) pour ce que constitue une identité d'utilisateur cohérente et utile pour l'organisation. Ces conditions incluent des attentes, implicites ou précédemment explicitées, sur la manière dont les données doivent se lier (un SSO = une personne, une adresse e-mail, etc.). Lorsque ces conditions ne sont pas remplies dans les données brutes (plusieurs e-mails pour un même SSO), l'agencement narratif

est mis à mal et un travail de *sensemaking* (Weick, 1995) est nécessaire pour le restaurer ou le redéfinir.

Dans ce processus, les données, une fois intégrées dans des réseaux de discours, de pratiques et de dispositifs, deviennent des actants capables d'influencer l'action (Latour, 2006). Elles participent à la structuration narrative de la situation dès leur apparition, en tant qu'éléments de mise en intrigue. Comme le souligne Cihodariu (2012), les récits ne sont *pas des reflets du réel*, mais des tentatives de « fixer le sens » dans un contexte où la vérité est toujours instable, construite et négociée. Le récit devient ainsi un dispositif épistémique, un acte de positionnement subjectif dans le monde social. Le Profil SSO, loin d'être un fait, une entité statique, est une biographie numérique de l'utilisateur, interprétée à travers des filtres techniques et organisationnels. Les choix de traitement de données – fusion, exclusion, archivage – traduisent des interprétations sur ce qui est (a)normal, pertinent ou stratégique. Les conventions de quantification et de contextualisation, si elles restent tacites, peuvent être imposées par un acteur dominant, faisant passer des jugements pour des faits(Colletis et al., 2020). Notre cas d'étude a mis en évidence une partie de cette « dispute » (Colletis et al., 2020, pp. 14-15)en montrant comment l'agencement des données génère des tensions – le « sourire crispé » de l'ingénieure face aux demandes du *Business* – qui révèlent les valeurs d'interopérabilité, de standardisation et de personnalisation à l'œuvre.

Ces résultats contribuent à nuancer la vision du « quatrième paradigme » (Hey et al., 2009; Kitchin, 2014)en montrant que la donnée est une « construction située » (Gitelman, 2013)et que son intelligibilité est toujours médiatisée par des interactions sociotechniques complexes. La donnée dispose d'une autonomie relative, constamment négociée et mise en forme par les acteurs humains, qui restent les « maîtres du sens » (Leleu-Merviel, 2003). Les « connaissances incarnées » de l'ingénieure de la donnée, c'est-à-dire sa connaissance des systèmes cibles et des conditions d'étude (Leonelli, 2014), sont essentielles pour interpréter les données et leur conférer une valeur probante.

Dans le contexte des organisations médiatiques, cette analyse met en lumière comment la performance chiffrée, incarnée par le Profil SSO, devient un levier de légitimation et de décision. Cependant, elle montre que cette légitimité n'est pas donnée par les chiffres seuls, mais par le travail invisible de qualification, de nettoyage et de mise en récit qui rend ces chiffres intelligibles et actionnables. Cela soulève des questions importantes sur l'équilibre entre la rationalisation algorithmique et la pluralité des récits et sur la manière dont les médias peuvent maintenir leur diversité éditoriale face à la pression des indicateurs (Burnier, 2023; Lambotte et al., 2018).

Cette étude, basée sur un cas d'étude unique, offre une connaissance *située* sur la co-constitution de l'identité de l'utilisateur sur les plateformes médiatiques. Cependant, la généralisation de ces résultats est limitée. De futures recherches pourraient explorer ces dynamiques dans d'autres contextes organisationnels ou avec d'autres types de

données massives. Par ailleurs, l'émergence de l'intelligence artificielle générative invite à explorer de nouvelles modalités d'agencement narratif et de ventriloquie. En manipulant des corrélations statistiques pour générer des contenus textuels ou visuels, ces outils opèrent une forme de mise en récit automatisée. Il s'agirait alors d'analyser comment ces systèmes, en tant qu'actants, reconfigurent la production de sens en stabilisant des récits par le calcul, tout en interrogeant la place de l'humain dans cette chaîne de ventriloquie machinique.

6. Références

Ananny, M., & Crawford, K. (2016). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media & Society, 20(3), 973-989. https://doi.org/10.1177/1461444816676645

Bachimont, B. (2014). Le nominalisme et la culture : questions posées par les enjeux du numérique. In B. Stiegler (Ed.), Digital Studies. Organologie des savoirs et technologies de la connaissance (pp. 63-78).

Bachimont, B. (2015). Le numérique comme milieu : enjeux épistémologiques et phénoménologiques. Principes pour une science des données. Interfaces numériques, 4. https://doi.org/10.3166/RIN.4.385-402

Bachimont, B. (2020). Rhétorique de l'intelligibilité, vers des humanités du numérique. MSH Sud, Montpellier.

Bencherki, N., Vásquez, C., Plourde, M.-C., & Brummans, B. H. J. M. (2023). Communiquer, c'est s'organiser. Une anthologie des écrits de James R. Taylor. Communiquer. Revue de communication sociale et publique(Communiquer, c'est s'organiser). https://doi.org/10.4000/communiquer.11086

Boyd, D., & Crawford, K. (2012). Critical Questions for Big Data. Provocations for a cultural, technological, and scholarly phenomenon. Information, Communication & Society, 15(5), 662-679. https://doi.org/10.1080/1369118x.2012.678878

Burnier, C. (2023). Fabriquer des données, produire du sens : une ethnographie dans les médias. Communication et organisation(64), 19- 31. https://doi.org/10.4000/communicationorganisation.12558

Cihodariu, M. (2012). Narratives as instrumental research and as attempts of fixing meaning1. The uses and misuses of the concept of "narratives". Journal of Comparative Research in Anthropology and Sociology, 3(2).

Colletis, G., Fieux, É., Isla, A., & Peneranda, A. (2020). De la donnée à l'information : de l'importance des conventions dans les organisations. Revue ouverte d'ingénierie des systèmes d'information, 1(2), 1-17.

Cooren, F. (2010). Ventriloquie, performativité et communication. Ou comment fait-on parler les choses. Réseaux, 5(163), 33-54. https://doi.org/10.3917/res.163.0033

Cooren, F. (2013). Manières de faire parler. Interaction et ventriloquie. Le Bord de l'eau.

Cooren, F., & Robichaud, D. (2006). Globaliser et disloquer en situation d'interaction: Comment asymétrise-t-on une relation ? Les interactions asymétriques. Québec: Editions Nota Bene, 113(32).

Couldry, N. (2017). The Myth of Big Data. In M. T. Schäfer & K. van Es (Eds.), The datafied society. Studying culture through data (pp. 235-239).
Gherardi, S. (2016). To start practice theorizing anew: The contribution of the concepts of agencement and formativeness. Organization, 23(5), 680- 698. https://doi.org/10.1177/1350508415605174
Gitelman, L. (2013). Raw data is an oxymoron. MIT press.
Hey, A. J. G., Tansley, S. W., & Tolle, K. M. (2009). The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft research.
Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. Big Data & Society, 1(1). https://doi.org/10.1177/2053951714528481
Lambotte, F., De Meyere, D., Andry, T., & Dumont de Chassart, P. (2018). L'intelligibilité des données massives de l'activité socionumérique : de la conception du dispositif à la construction de sens en contexte. Communication et organisation, 54, 149- 167. https://doi.org/10.4000/communicationorganisation.7054
Latour, B. (2006). Changer de société. Refaire de la sociologie. Paris, La Découverte.
Leleu-Merviel, S. (2003). Les désarrois des "Maîtres du sens" à l'ère du numérique. H2PTM'03, Laboratoire des Sciences de la Communication, de l'Université de Valenciennes et du Hainaut-Cambrésis.
Leonelli, S. (2014). Data Interpretation in the Digital Age. Perspectives on Science, 22(3), 397-417. https://doi.org/10.1162/POSC_a_00140
Marion, P. (1997). Narratologie médiatique et médiagénie des récits. Recherches en communication(7), 61-87.
Mayer-Schönberger, V., & Cukier, K. (2013). Big Data: A Revolution That Will Transform How We Live, Work, and Think. Massachusetts: Houghton Mifflin Harcourt.
Nathues, E., van Vuuren, M., & Cooren, F. (2020). Speaking about vision, talking in the name of so much more: A methodological framework for ventriloquial analyses in organization studies. Organization Studies, 42(9), 1457- 1476. https://doi.org/10.1177/0170840620934063
Ricoeur, P. (2014). La Mémoire, l'Histoire, l'Oubli. Le Seuil. https://doi.org/10.3917/ls.ricoe.2014.05
Rodrigues, P. (2019). « Empêtrés dans des histoires ». Mélanges de science religieuse, 76(3), 55-65.
Rouleau, L. (2013). L'ethnographie organisationnelle d'hier à demain. Revue internationale de psychosociologie et de gestion des comportements organisationnels, 19(48), 27-43.
Schoeneborn, D., Vásquez, C., & Kuhn, T. (2025). Communication-as-Constitutive Perspectives on Organization. Academy of Management Collections, 4(1), 1- 19. https://doi.org/10.5465/amc.2023.0007
Schulz, M. (2017, 30 mai-2 juil 2017). Des conditions d'émergence du sens aux conditions d'instauration des discours. Greimas aujourd'hui : l'avenir de la structure. Actes du congrès de l'Association Française de Sémiotique, Unesco.
Taylor, J. R. (1988). Une organisation n'est qu'un tissu de communications. Department of Communication, University of Montreal.
Taylor, J. R., Cooren, F., Giroux, H., & Robichaud, D. (1996). The Communicational Basis of Organization: Between the Conversation and the Text. Communication Theory, 6(1),

1- 39.
van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. Surveillance & Society, 12(2), 197-208. Van Maanen, J. (2011). Ethnography as Work: Some Rules of Engagement. Journal of Management Studies, 48(1), 218-234. https://doi.org/10.1111/j.1467-6486.2010.00980.x
Weick, K. E. (1995). Sensemaking in organizations (Vol. 3). Sage.
Weick, K. E. (2020). Sensemaking, Organizing, and Surpassing: A Handoff. Journal of Management Studies, 57(7), 1420-1431. https://doi.org/10.1111/joms.12617

7. Notes

[1] Toutes les traductions de langue anglaise présentes dans cet article sont de l'autrice.
[2] Le terme métier ou Business, omniprésent dans les organisations observées, désigne les acteurs dont les activités sont directement liées à la mission principale de l'entreprise (rédaction, publicité, marketing). Il se distingue des fonctions supports (informatique, ressources humaines) qui soutiennent l'activité sans constituer le cœur du métier.